# Efficiency of a class of unbiased estimators for the invariant distribution function of a diffusion process [*]

Ilia Negri[†]

## Abstract

We consider the problem of the estimation of the invariant distribution function of an ergodic diffusion process when the drift coefficient is unknown. The empirical distribution function is a natural estimator which is unbiased, uniformly consistent and efficient in different metrics. Here we study the properties of optimality for an other kind of estimator recently proposed. We consider a class of unbiased estimators and we show that they are also efficient in the sense that their asymptotic risk, defined as the integrated mean square error, attains an asymptotic minimax lower bound.

**Key words:** ergodic diffusion, asymptotically efficient estimators, lower bound.
**2000 MSC:** 60G35; 62M20.

---

[†]Department of Management and Information Technology, Viale Marconi 5, 24044 Dalmine (BG), Italy. `ilia.negri@unibg.it`

# 1   Introduction

We consider the problem of the estimation of the distribution function $F(x)$, $x \in \mathbb{R}$ by the observation of a diffusion process $\{X_t : 0 \leq t \leq T\}$. We suppose that the process $X_t$, $t \geq 0$ possesses the ergodic property with invariant measure $\mu$ and $F(x) = \mu\{(-\infty, x]\}$. A natural estimator for $F(x)$, $x \in \mathbb{R}$, is the empirical distribution function $\hat{F}_T(x)$. It is well known that this estimator is uniformly consistent by the Glivenko–Cantelli theorem and asymptotically normal (Kutoyants, 1997). The problem of the asymptotically efficiency of the empirical distribution function has been considered for different model and different metrics. For the model of independent and identical distributed random variables the empirical distribution function is asymptotically efficient, in a global framework, in the sense that its integrated mean square error attains the lower bound given for all the estimators of the distribution function. Such result has been established earlier by Levit (1978) and Millar (1979) using the theory of local asymptotic normality. Gill and Levit (1995) obtained the same result using a different approach based on a multidimensional version of the van Trees inequality. The same approach introduced by Gill and Levit was successfully applied in Kutoyants and Negri, (2001) to prove that the empirical distribution function is asymptotically efficient in the problem of invariant distribution estimation for ergodic diffusion processes. For the same model Negri (1998) has proved the asymptotically efficiency of the empirical distribution function when the metric utilized in the risk function is based on the *sup* norm.

Recently (Kutoyants, 2004) a class of unbiased estimator for the invariant distribution function has been introduced. These estimators, that do not contain the empirical distribution function as particular case, are consistent and asymptotically normal. In this work we prove that they are also asymptotically efficient in the sense that their integrated mean square error attain the lower bound given for all the estimators of the invariant distribution function. As in the case of the estimation of the invariant density, we have many efficient estimators, so the problem of finding the second order efficient estimator arise naturally (see Dalalyan and Kutoyants 2004 where the problem is considered for the invariant density estimation). This problem it is not considered here, but it will be an

argument of future researches.

The note is organized as follow. In the next section we present the statement of the problem and the assumptions. In Section 3 we present the lower bound for the risk. In Section 4 we prove that the class of unbiased estimator attain this bound and finally, in Section 5, we give same examples of such estimators.

## 2 Preliminaries

In this section we introduce the model and its first properties, while the statistical problem will be presented in the next section. Let us consider a one dimensional diffusion process

$$\mathrm{d}X_t = S(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0, \quad t \geq 0 \tag{1}$$

where $\{W_t : t \geq 0\}$ is a standard Wiener process, and the initial value $X_0$ is independent of $W_t$, $t \geq 0$. The drift coefficient $S$ will be supposed unknown to the observer and the diffusion coefficient $\sigma^2$ will be a known positive function. Let us introduce the condition:

$\mathcal{ES}$. *The function $S$ is locally bounded, the function $\sigma^2$ is positive and continuous and for some $A > 0$ the condition $xS(x) + \sigma(x)^2 \leq A(1 + x^2)$, $x \in \mathbb{R}$ holds.*

Under the condition $\mathcal{ES}$ the equation (1) has an unique weak solution (see Durrett, 1996, p. 210). To guarantee ergodicity we introduce the following condition:

$\mathcal{RP}$. *The function $S$ and $\sigma$ are such that:*

$$V_S(x) = \int_0^x \exp\left\{-2\int_0^y \frac{S(v)}{\sigma(v)^2}\mathrm{d}v\right\}\mathrm{d}y \to \pm\infty, \quad as \quad x \to \pm\infty$$

*and*

$$G(S) = \int_{-\infty}^{+\infty} \frac{1}{\sigma(x)^2}\exp\left\{2\int_0^x \frac{S(v)}{\sigma(v)^2}\mathrm{d}v\right\}\mathrm{d}x < +\infty.$$

If the condition $\mathcal{RP}$ is satisfied then the weak solution of (1), $\{X_t, t \geq 0\}$, has the ergodic property (see for example Gikhman and Skorohod, 1972), that is, there exists an invariant probability measure $\mu_S$ such that for every measurable function $g$ such that $\mathbf{E}_S|g(\xi)| < \infty$, we have with probability one,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T g(X_t)\mathrm{d}t = \int_{\mathbb{R}} g(z)f_S(z)\mathrm{d}z = \mathbf{E}_S(g(\xi))$$

where $\xi$ has the invariant measure as distribution, $\mathbf{E}_S$ denote the mathematical expectation with respect to $\mu_S$, and $f_S$ is the invariant density given by

$$f_S(y) = \frac{1}{G(S)\sigma(y)^2} \exp\left\{2 \int_0^y \frac{S(v)}{\sigma(v)^2}\mathrm{d}v\right\}. \tag{2}$$

Suppose we observe different diffusion processes $\{X_t : 0 \le t \le T\}$ given by equation (1) with drift coefficients respectively given by $S_1$, $S_2$ and $S_0 = 0$ and initial value respectively $X_0^1$, $X_0^2$ and $X_0^0$. Let us introduce the following condition.

$\mathcal{EM}$. *The functions $S_1$, $S_2$ and $\sigma$ satisfy condition $\mathcal{ES}$ and the densities (with respect to the Lebesgue measure) of the corresponding initial values $X_0^1$, $X_0^2$ and $X_0^0$ have the same support (if the initial value is nonrandom, then we suppose that it takes the same value for all processes).*

If condition $\mathcal{EM}$ holds true, all the measures $P_S^T$, for different $S$, induced by the process $\{X_t, : 0 \le t \le T\}$, in the space $\mathcal{C}_T$, the space of all the continuos function on $[0, T]$ with uniform metric and Borel $\sigma$-algebra $\mathcal{B}(\mathcal{C}_T)$, are equivalent. See Kutoyants, 2004, p. 34 and the references therein.

# 3   The asymptotic global bound

Given the diffusion process (1) we suppose that conditions $\mathcal{ES}$ and $\mathcal{RP}$ are satisfied, that $X_0$ has density $f_S$, given by (2), so the process $\{X_t, t \ge 0\}$ is ergodic and strictly stationary. We are interested in the estimation of the invariant distribution function

$$F_S(x) = \int_{-\infty}^x \frac{1}{G(S)\sigma(y)^2} \exp\left\{2 \int_0^y \frac{S(v)}{\sigma(v)^2}\mathrm{d}v\right\} \mathrm{d}y \tag{3}$$

by the observation $X^T = \{X_t : 0 \le t \le T\}$ solution of (1) when $\sigma$ is known and $S$ is unknown. Let us denote by $\mathbf{E}_S^T$ the mathematical expectation with respect to the measure $P_S^T$.

Let $\bar{F}_T(x)$ be any estimator of (3) for $x \in \mathbb{R}$. We define the integrated mean square error as

$$\rho_T(\bar{F}_T, F_S) = T\mathbf{E}_S^T \int_{\mathbb{R}} |\bar{F}_T(x) - F_S(x)|^2 \nu(\mathrm{d}x) \tag{4}$$

4

where $\nu$ is a finite measure on $\mathbb{R}$. We will refer also to it as *global risk*. A natural estimator of $F_S(x)$ for $x \in \mathbb{R}$ is the empirical distribution function defined as follows

$$\hat{F}_T(x) = \frac{1}{T} \int_0^T \chi_{\{X_t < x\}} \mathrm{d}t, \quad x \in \mathbb{R}. \tag{5}$$

This estimator is uniformly consistent, asymptotically normal and asymptotically efficient in the sense that the empirical distribution function achieves a local asymptotic minimax lower bound for the integrated mean square error of an arbitrary estimator. For a fixed function $\sigma$ let us introduce the classes $\mathcal{S}_\sigma = \{S : \text{conditions } \mathcal{ES}, \mathcal{EM}, \mathcal{RP} \text{ are fulfilled}\}$ and $\mathcal{S}_\sigma^* \subset \mathcal{S}_\sigma$ such that for every $S_*$ in $\mathcal{S}_\sigma^*$ there exist a $\delta > 0$, and a vicinity $V_\delta = \{S : \sup_{x \in \mathbb{R}} |S_*(x) - S(x)| < \delta, \quad S \in \mathcal{S}_\sigma^*\}$, such that $\sup_{S \in V_\delta} G(S) < +\infty$.

For $x$ and $y$ in $\mathbb{R}$ we denote by $x \wedge y$ and by $x \vee y$ respectively the minimum and the maximum between $x$ and $y$. Let us introduce the function

$$R_S(x, y) = 4 \int_{-\infty}^{+\infty} \frac{F_S(v \wedge x)\big(1 - F_S(v \vee x)\big) F_S(v \wedge y)\big(1 - F_S(v \vee y)\big)}{\sigma(v)^2 f_S(v)} \mathrm{d}v.$$

and the quantity

$$\rho_*(S) = \int_{\mathbb{R}} 4\mathbf{E}_S \left( \frac{F_S(x) F_S(\xi) - F_S(\xi \wedge x)}{\sigma(\xi) f_S(\xi)} \right)^2 \nu(\mathrm{d}x) = \int_{\mathbb{R}} R_S(x, x) \nu(\mathrm{d}x)$$

Let us introduce the following condition.

$\mathcal{Q}_1$. *The function $S_* \in \mathcal{S}_\sigma^*$ and for some $\delta > 0$*

$$\sup_{S \in V_\delta} \rho_*(S) = \sup_{S \in V_\delta} \int_{\mathbb{R}} 4\mathbf{E}_S \left( \frac{F_S(x) F_S(\xi) - F_S(\xi \wedge x)}{\sigma(\xi) f_S(\xi)} \right)^2 \nu(\mathrm{d}x) < +\infty.$$

We have the following result (Kutoyants and Negri, 2001).

**Theorem 1.** *Let $S_* \in \mathcal{S}_\sigma^*$ and condition $\mathcal{Q}_1$ be fulfilled. Then*

$$\lim_{\delta \to 0} \lim_{T \to \infty} \inf_{\bar{F}_T} \sup_{S \in V_\delta} \rho_T(\bar{F}_T, F_S) \geq \rho_*(S^*)$$

*where the* inf *is taken over any estimator $\bar{F}_T$ of $F_S$.*

The definition of asymptotically efficient estimator arise naturally from the above theorem.

**Definition 1.** *Let condition $\mathcal{Q}_1$ be fulfilled. Then an estimator $\hat{F}_T$ is called asymptotically efficient if for any $S_* \in \mathcal{S}_\sigma^*$ we have*

$$\lim_{\delta \to 0} \lim_{T \to \infty} \sup_{S \in V_\delta} \rho_T(\hat{F}_T, F_S) = \rho_*(S^*) \tag{6}$$

Put

$$H_{x,S}(y) = 2 \int_0^y \frac{F_S(v \wedge x) - F_S(v)F_S(x)}{\sigma(v)^2 f_S(v)} \mathrm{d}v. \tag{7}$$

Let us introduce the following condition.

$\mathcal{Q}_2$. *Suppose that*

$$\sup_{S \in V_\delta} \int_{\mathbb{R}} \mathbf{E}_S H_{x,S}(\xi)^2 \nu(\mathrm{d}x) < +\infty.$$

The following theorem establishes the asymptotic efficiency of the empirical distribution function (5) in the sense of equality (6).

**Theorem 2.** *Let conditions $\mathcal{Q}_1$, $\mathcal{Q}_2$ hold and $\rho_*(S)$ be continuous in the uniform topology at the point $S^*$, then the empirical distribution function is asymptotically efficient.*

It is proved in Kutoyants and Negri (2001).

# 4   A class of unbiased estimators

In this section we consider a class of estimators of $F_S(x)$ recently introduced in Kutoyants (2004) defined, for $x \in \mathbb{R}$ as

$$\tilde{F}_T(x) = \frac{1}{T} \int_0^T R_x(X_t)\mathrm{d}X_t + \frac{1}{T} \int_0^T N_x(X_t)\mathrm{d}t \tag{8}$$

where $R_x(y) = 2\chi_{\{y<x\}}K_x(y)h(y)$, $N_x(y) = \chi_{\{y<x\}}K_x(y)h'(y)\sigma^2(y)$, $K_x(y) = \int_y^x \frac{\mathrm{d}v}{\sigma^2(v)h(v)}$ and $h$ is a positive and continuously differentiable function. It can be proved that these estimators, for different functions $R_x$ and $N_x$ are all unbiased, consistent and asymptotically normal for a fixed $x$. Let us suppose the following conditions hold:

$$\mathbf{E}_S(R_x(\xi)\sigma(\xi))^2 < +\infty, \quad \mathbf{E}_S|N_x(\xi)| < +\infty, \quad \lim_{y \to -\infty} R_x(y)\sigma(y)^2 f_S(y) = 0 \tag{9}$$

We have the following result (Kutoyants, 2004).

**Theorem 3.** *Let $S \in \mathcal{S}_\sigma$, $R_S(x,x) < +\infty$, and conditions (9) be fulfilled. Then the estimator $\tilde{F}_T(x)$ is unbiased, consistent and asymptotically normal with variance given by $R_S(x,x)$.*

Let us define the following function

$$G_x(y) = 2 \int_0^y \chi_{\{v<x\}} K_x(v) h(v) \mathrm{d}v \tag{10}$$

We introduce the following condition.

$\mathcal{Q}_3$. *Suppose that*

$$\sup_{S \in V_\delta} \int_{\mathbb{R}} \mathbf{E}_S G_x(\xi)^2 \nu(\mathrm{d}x) < +\infty.$$

The following theorem proves that the new class of estimators defined by (8) is asymptotically efficient.

**Theorem 4.** *Let conditions $\mathcal{Q}_1$, $\mathcal{Q}_2$ and $\mathcal{Q}_3$ hold true. Let $\rho_*(S)$ be continuous in the uniform topology at the point $S^*$, then*

$$\lim_{\delta \to 0} \lim_{T \to \infty} \sup_{S \in V_\delta} \rho_T(\tilde{F}_T, F_S) = \rho_*(S^*) \tag{11}$$

*Proof.* Let us define $c_x(y) = \chi_{\{y<x\}} K_x(y)[2h(y)S(y) + h'(y)\sigma^2(y)] - F_S(x)$ and $d_x(y) = 2\chi_{\{y<x\}} h(y) K_x(y)\sigma(y)$. We have the following representation of the empirical process

$$\sqrt{T}\left(\tilde{F}_T(x) - F_S(x)\right) = \frac{1}{\sqrt{T}} \int_0^T c_x(X_t)\mathrm{d}t + \frac{1}{\sqrt{T}} \int_0^T d_x(X_t)\mathrm{d}W_t. \tag{12}$$

Now we search for a function $M_{x,S}$ such that

$$M'_{x,S}(y)S(y) + \frac{1}{2}M''_{x,S}(y)\sigma^2(y) = c_x(y). \tag{13}$$

Putting $M'_{x,S} = m$, the equation (13) can be rewritten as $m' = \frac{2c}{\sigma^2} - \frac{2S}{\sigma^2}m$ which have solution

$$m(z) = \frac{2}{f(z)\sigma^2(z)} \int_{-\infty}^z c_x(v) f(v)\mathrm{d}v. \tag{14}$$

Integrating by part the integral in (14) and observing that

$$\frac{\mathrm{d}}{\mathrm{d}y}\sigma^2(y)f(y) = 2S(y)f(y)$$

the function $m$ can be rewritten as

$$m(z) = 2\chi_{\{z<x\}} h(z) K_x(z) + 2\frac{F(x \wedge z) - F(x)F(z)}{\sigma^2(z)f(z)}. \tag{15}$$

7

Choosing the function $M_{x,S}$ such that $M_{x,S}(0) = 0$, it has this form

$$M_{x,S}(y) = \int_0^y 2\chi_{\{z<x\}}h(z)K_x(z)\mathrm{d}z + \int_0^y 2\frac{F(x\wedge z) - F(x)F(z)}{\sigma^2(z)f(z)}\mathrm{d}z \tag{16}$$

From (7), (10) and (16) we can write

$$M_{x,S}(y) = G_x(y) + H_{x,S}(y). \tag{17}$$

By the Itô formula we can write

$$\mathrm{d}M_{x,S}(X_t) = \left(M'_{x,S}(X_t)S(X_t) + \frac{1}{2}M''_{x,S}(X_t)\sigma^2(X_t)\right)\mathrm{d}t + M'_{x,S}(X_t)\sigma(X_t)\mathrm{d}W_t.$$

Now we can substitute the Lebesgue integral in (12) by means of this formula and the empirical process (12) became

$$\sqrt{T}\left(\tilde{F}_T(x) - F_S(x)\right) = \frac{M_{x,S}(X_T) - M_{x,S}(X_0)}{\sqrt{T}} + \frac{1}{\sqrt{T}}\int_0^T 2\frac{F_S(x\wedge X_t) - F_S(x)F_S(X_t)}{\sigma(X_t)f_S(X_t)}\mathrm{d}W_t.$$

From (17) and conditions $\mathcal{Q}_2$ and $\mathcal{Q}_3$ it follows that

$$\lim_{T\to+\infty}\sup_{S\in V_\delta}\int_{\mathbb{R}}\mathbf{E}_S\left(\frac{M_{x,S}(X_T) - M_{x,S}(X_0)}{\sqrt{T}}\right)^2\nu(\mathrm{d}x) = 0 \tag{18}$$

Moreover by the continuity of $\rho_*(S)$ at the point $S^*$, as in Kutoyants and Negri (2001) we can conclude

$$\lim_{\delta\to 0}\lim_{T\to+\infty}\sup_{S\in V_\delta}\int_{\mathbb{R}}\mathbf{E}_S\left(\frac{1}{\sqrt{T}}\int_0^T 2\frac{F_S(x\wedge X_t) - F_S(x)F_S(X_t)}{\sigma(X_t)f_S(X_t)}\mathrm{d}W_t\right)^2\nu(\mathrm{d}x) = \rho_*(S^*) \tag{19}$$

So by (18) and (19) it follows (11), and the proof is concluded. $\qquad\square$

## 5   Examples

In Theorem 3 and Theorem 4, conditions $\mathcal{Q}_1$ and $\mathcal{Q}_2$ involve the function $S$, indeed the model. Example of function $S$ for which conditions $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are satisfied can be found in Kutoyants and Negri (2001) and in Kutoyants (2004). Conditions (9) in Theorem 3 and $\mathcal{Q}_3$ in Theorem 4 are on the class of estimators. The class of unbiased estimators defined by (8) is very general and is not empty. In this section we will show that for a very large choice of functions $h(\cdot)$ the related estimator $\tilde{F}$ satisfy condition $\mathcal{Q}_3$ and conditions (9). In all this section let us consider for simplicity $\sigma(u) = 1$.

**Example 1.** Let us consider $h(u) = 1 + u^{2p}$, $p \geq 1$. We have

$$K_x(y) = \int_v^x \frac{du}{1 + u^{2p}} = \tilde{K}_p(x) - \tilde{K}_p(y) \tag{20}$$

where $\tilde{K}_p(\cdot)$ is a primitive of $1/h(\cdot)$. We have that for every $p \geq 1$, $|K_x(y)| \leq \pi$. From (20) we have $|R_x(y)| = |2\chi_{\{y<x\}} K_x(y) h(y)| \leq 2\pi |h(y)|$ and $\mathbf{E}_S(R_x(\xi))^2 \leq 2\pi \int_{\mathbb{R}} (1 + y^{2p}) f_S(y) dy < +\infty$ because the invariant distribution admits the moment of every order. Moreover $N_x(y) = \chi_{\{y<x\}} K_x(y) 2py^{2p-1}$ and we have $\mathbf{E}_S(|N_x(\xi)|) \leq +\infty$. Finally $\lim_{y \to -\infty} R_x(y) f_S(y) = 0$ and conditions (9) are satisfied. Let now consider $G_x(y) = \int_0^y R_x(v) dv$. We have

$$\mathbf{E}_s(G_x(\xi)^2) = \int_{-\infty}^{+\infty} \left( \int_0^y R_x(v) dv \right)^2 f_S(y) dy < C \tag{21}$$

where $C$ is a constant that does not depend on $S \in V_\delta$. Equation (21) imply that condition $\mathcal{Q}_3$ is satisfied without any further condition on measure $\nu$. Note that for $p = 1$, the (20) became $K_x(v) = \int_v^x \frac{du}{1+u^2} = \text{artg}x - \text{artg}v \leq \pi$ for every $v$ and $x$.

**Example 2.** Let us consider $h(u) = e^{\delta u}$, $\delta > 0$. We have

$$K_x(y) = \int_y^x \frac{du}{e^{\delta u}} = \frac{e^{-\delta y} - e^{-\delta x}}{\delta}$$

and

$$R_x(y) = -2\chi_{\{y<x\}} \frac{e^{\delta(y-x)}}{\delta}. \tag{22}$$

From (22) it follows that $R_x(y)$ is bounded with respect to both the variables $x$ and $y$. In virtue of this fact conditions (9) and $\mathcal{Q}_3$ can easily be check. Also in this case we have not to set any further condition on measure $\nu$.

**Example 3.** Let us consider $h(u) = c$, where $c$ is a real constant. We have $K_x(v) = \frac{x-v}{c}$, $R_x(v) = 2\chi_{v<x}(x - v)$ and $N_x(v) = 0$. In this case the first of the conditions (9) is not satisfied. In fact it follows that $\mathbf{E}_S(R_x(\xi))^2 = 4x^2 F_S(x) + C$, where $C$ is a constant depending on the moment values of the invariant distribution. For each fixed $x$, $\mathbf{E}_S(R_x(\xi))^2$ is finite but when $x \to +\infty$, also $\mathbf{E}_S(R_x(\xi))^2$ goes to infinity. In any case we have $G_x(y) = 2x(y \wedge x) - (y \wedge x)^2$ and $\int_{\mathbb{R}} \mathbf{E}_S(G_x(\xi)^2) \nu(dx)$ is finite if the measure $\nu$ admits the moment of order fourth.

We observe that the empirical distribution function does not belong to this class of estimators. Indeed it would be necessary that $R_x(y) = 0$ and $K_x(y)h(y) = 1$ for every $y$ and $x$ belonging to $\mathcal{R}$. But if such a function $h$ exists it has to depend on $x$ and this is not possible.

## Acknowledgments

# References

Dalalyan, A.S., Kutoyants, Y.A. (2004), On second order minimax estimation of invariant density for ergodic diffusion. *Statist. Decisions*, **22**, 1, 17–41.

Durrett, R., (1996), *Stochastic Calculus. A Pratical Introduction*, (CRC Press, New York).

Gikhman, I.I., Skorohod, A.V. (1972), *Stochastic Differential Equations*, (Springer–Verlag, New York).

Gill, R.D., Levit, B.Y. (1995), Applications of the van Trees inequality: a Bayesian Cramér-Rao bound, *Bernoulli*, **1+2**, 59-79.

Kutoyants, Y.A. (1997), Efficiency of the empirical distribution for ergodic diffusion, *Bernoulli*, **3**, 4, 445-456.

Kutoyants, Y.A. (2004), *Statistical Inference for Ergodic Diffusion Processes*, (Springer, New York).

Kutoyants Y.A., Negri I. (2001), On $L_2$-efficiency of empiric distribution for diffusion process, *Theory of Probability and its Applications*, **46**, 1, 164-169.

Levit, B.Y. (1978), Infinite-dimensional information inequalities, *Theory Prob. Appl.*, **23**, 371-377.

Millar, P.W. (1979), Asymptotic minimax theorems for the sample distribution function, *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **48**, 233-252.

Negri, I. (1998), Stationary distribution function estimation for ergodic diffusion process, *Statistical Inference for Stochastic Processes*, **1**, 61–84.